

A Bottom Line for the LHC Data by Leveraging Pileup as a Zero Bias Trigger

Benjamin Nachman*

Lawrence Berkeley National Laboratory

Francesco Rubbo†

SLAC National Accelerator Laboratory

(Dated: March 24, 2017)

Due to a limited bandwidth and a large proton-proton interaction cross-section relative to the rate of interesting physics processes, most events produced at the Large Hadron Collider (LHC) are discarded in real time. A sophisticated *trigger* system must quickly decide which events should be kept and is very efficient for a broad range of processes. However, there are many processes that cannot be accommodated by this trigger system. Furthermore, there may be models of physics beyond the Standard Model (BSM) constructed after data taking that could have been triggered, but no trigger was implemented at run time. Both of these cases can be covered by exploiting pileup interactions as an effective zero bias trigger. After the LHC, this zero bias stream will have accumulated about 1 fb^{-1} of data from which a bottom line cross-section limit of $\mathcal{O}(1) \text{ fb}$ can be set for BSM models already in the literature and those yet to come.

I. INTRODUCTION

At a proton-proton (pp) collider like the Large Hadron Collider (LHC), interesting events are rare. Unlike electron-positron colliders, the partonic center-of-mass energy \sqrt{s} follows a broad distribution set by parton distribution functions (PDF). As such, the total inelastic cross-section ($\mathcal{O}(100) \text{ mb}$ [1, 2] at $\sqrt{s} = 13 \text{ TeV}$) is many orders of magnitude above the production cross-section for electroweak scale particles, such as the W boson ($\mathcal{O}(10) \text{ nb}$ [3, 4] at $\sqrt{s} = 13 \text{ TeV}$). In order to increase the rate of interesting events as much as possible, the LHC is operated at very high luminosities. At its design luminosity, proton bunches collide every 25 ns, and the bunch density is such that multiple pp interactions (*pileup*) occur in each bunch crossing. Due to limited readout and disk space capabilities, it is not possible to fully record every bunch collision. Therefore, both the ATLAS and CMS experiments have developed strategies to *trigger* on events of interest. Trigger systems are implemented at multiple levels, with ultra-fast but simple algorithms in hardware (L1) and increasingly complex algorithms in software at higher levels (HLT), where a more detailed readout of the detectors is exploited. An event is fully recorded only if it satisfies the selection criteria at all levels of the trigger. The L1 triggers decrease the 40 MHz rate down to $\mathcal{O}(100) \text{ kHz}$, which is further reduced to $\mathcal{O}(1) \text{ kHz}$ after the HLT triggers. In order to achieve these reductions, triggers targeting processes with a very high cross-section are *prescaled*: events are randomly discarded so that only a fraction $1/p$ ($p = \text{prescale}$) are recorded. For example, at $\sqrt{s} = 8 \text{ TeV}$, the ATLAS single jet triggers targeting events with at least one jet with a minimum transverse momentum (p_T) between 50 and

100 GeV had prescales of $p \sim 10^4$ [5]. The lowest unprescaled ($p = 1$) single jet trigger requires a minimum transverse momentum $p_T \sim 500 \text{ GeV}$. The prescales for low p_T processes, such as inclusive jet production, are increased linearly with the luminosity, in order to keep the rate constant.

While the existing trigger system is very effective at identifying high p_T objects, there are a plethora of viable models of physics beyond the Standard Model (BSM) that are not well-covered. One broad class of models predicts exotic signatures involving isolated charged particle tracks. Pattern recognition in the ATLAS and CMS inner detectors is computationally expensive and it only runs in a limited way at the HLT (and possibly at L1 in the future). Tracks with kinks, displaced vertices, high dE/dx , anomalous timing, intermittent hits, and exotic curvature will not be covered by L1 tracking and are also not easily (or at all) covered in the HLT (see Ref. [6] for a review). For example, oscillating pairs of tracks from new strong dynamics [7–9] require dedicated reconstruction algorithms. In addition to models with low multiplicity tracks, BSM processes that predict extreme multiplicities of low energy particles [9–11] are striking signatures that may be largely uncovered by existing or even possible triggering techniques. Another broad class of models predicts exotic structure inside hadronic jets. This includes jets with many displaced vertices [12] as well as jets with large invisible components [13]. There are likely many other models that have yet to be proposed in the literature that would leave extraordinary detector signatures but too exotic to be captured by standard trigger schemes.

All of the models discussed so far have the property that their signature is so exotic that there is likely not a significant background rate from the Standard Model. The current triggering scheme also misses models that have a large SM background, such that they receive a large prescale. This includes low mass dijet resonances, such as leptophobic Z' [14].

*Electronic address: bnachman@cern.ch

†Electronic address: rubbo@cern.ch

There are two existing strategies for recovering model coverage that would otherwise have been lost by the trigger. One strategy is to look for a target process produced in association with another very energetic object that can be used for triggering. For example, a low mass Z' that decays into jets can be produced in association with a high p_T photon [15] or jet [16] from Initial State Radiation (ISR). However, this strategy introduces a large effective prescale due to a reduction in the cross-section. In addition, this procedure cannot be used to measure new or Standard Model processes at low p_T . Another powerful strategy, referred to as data-scouting or trigger-level analysis, stores only a smaller relevant fraction of the detector information for the events selected by the L1 trigger [17, 18]. These trigger-level analyses are not impacted by the prescales of the HLT triggers, but are limited by the L1 prescales that are often tighter. As only a small fraction of the detector information can be recorded at the L1 accept rate, only the specific final states for which the trigger-level analysis strategy has been designed are accessible.

The new strategy presented in this paper uses each individual pileup interaction for physics analysis. All of these interactions potentially contain interesting physics processes and are recorded by the detector along with the primary interaction that satisfied any arbitrary trigger. Every event passing any trigger can be used for the purpose of studying the pileup interactions, which are recorded with nearly no selection bias. The effective prescale associated to this *Zero Bias Trigger* (ZBT) is proportional to the overall trigger bandwidth. In addition, a trigger-level analysis can be combined with this strategy, thus enhancing even further the physics reach by enabling access to a large quantity of otherwise unused data.

II. THE ZERO BIAS TRIGGER

Reconstructed tracks from charged particles are the most important handle for identifying pileup interactions. Individual collision vertices are built from tracks and various objects can be associated to these vertices through their associated tracks. For example, the jet vertex tagger (JVT) used in ATLAS is 90% efficient at associating a jet with $20 < p_T < 50$ GeV to its correct vertex while mis-identifying stochastic or QCD jets from other vertices 1% of the time [19]. Ignoring the small detector inefficiencies and fake rates, the effective luminosity from the pileup collisions collected from a trigger system with bandwidth w is given by

$$\mathcal{L}(\text{ZBT}) = \frac{w}{40 \text{ MHz}} \times \int \mathcal{L} dt. \quad (1)$$

The last term in Eq. 1 is the integrated luminosity col-

lected with standard triggers. One way to derive this equation is to consider the number of events recorded by the LHC experiments. Suppose there is a process X with a cross-section σ_X . If Σ is the total inelastic cross-section, then a fraction σ_X/Σ (on average) of all pp collisions recorded will contain the process X . If the bandwidth is w and assuming that the production of X does not significantly influence this rate, then the total number of X events will be $\langle \mu \rangle \times w \times T \times \sigma_X/\Sigma$, where T is the amount of time the LHC is operated for pp collisions and $\langle \mu \rangle$ is the average number of pp collisions per bunch crossing. Now let Y be any Standard Model process that can be triggered with 100% efficiency and has a prescale of 1. Analogously to the calculation for X , the number of Y events recorded will be $\langle \mu \rangle \times H \times T \times \sigma_Y/\Sigma$, where H is the total rate of bunch crossings (40 MHz at the LHC). This shows that $\int \mathcal{L} dt = \langle \mu \rangle \times H \times T \times \Sigma$ and substituting this in for the equation for the number of X events and solving for the effective luminosity results in Eq. 1. The bandwidth w can go from the 100-500 Hz typical of Run 1 data-taking to the upper limit of the expected L1 trigger bandwidth for the High Luminosity LHC (HL-LHC) experiments [20, 21]. Table I shows the ZBT luminosity for various LHC running conditions in Run 1 and projected for Run 2 and beyond. The effective prescale for the best HL-LHC data acquisition scenarios (last row in Table I) is between 4000 and 6000. For trigger-level analyses of the ZBT dataset at the HL-LHC, the effective prescale is between 50 and 100. This means that if a particular signature has a primary trigger efficiency that is less than 0.02-0.03% offline or 1-2% at L1, then the ZBT dataset will record more signal events. Cross-section limit projections are described in the next section.

LHC Run	Total Lumi. [1/fb]	$\langle \mu \rangle$	L1 Rate [MHz]	HLT Rate [kHz]	ZBT [1/fb]	ZBT @ HLT [1/fb]
1	20	20	0.1	0.1	5×10^{-5}	0.05
2+3	300	80	0.1	1	7.5×10^{-3}	0.75
4+5 (ATLAS)	3000	200	0.4	10	0.75	30
4+5 (CMS)	3000	200	0.75	7.5	0.56	56.3

TABLE I: The ZBT luminosity for various LHC running conditions in Run 1 and projected for Run 2 and beyond. The last column shows the projected luminosity when performing a trigger-level analysis, before the application of any software prescales.

III. CROSS-SECTION PROJECTIONS

Figure 1 presents the 95% confidence level cross-section limits with the CL_s procedure [22], covering a broad class of models, for the ZBT throughout the lifetime of the

(HL-)LHC. Limits are computed assuming various signal-to-background ratios. For a zero-background search, at least three signal events are needed to reach the 95% confidence limit. The predicted integrated luminosity during Runs 2 and 3 of the LHC is about 300 fb^{-1} while the final HL-LHC data is expected to be about 3 ab^{-1} .

The offline version of the ZBT can be analyzed at any time, including (well) after the full HL-LHC program has ended, while the trigger-level (online) version requires some level of analysis to be implemented in the software trigger at run time. For a model that predicts an extraordinary signature with nearly no SM background (see Sec. I), a ZBT analysis will be able to set a cross-section limit of nearly 3 fb^{-1} . For reference, this is the cross-section of a 1.8 TeV gluino [23], a 1.1 TeV stop [23], a 1.4 TeV fermionic top quark partner [24], and a 1.7 TeV colored quirk with infracolor representation size 5 [9]. Cross-section limits for the ZBT applied at HLT are nearly a factor of 100 stronger, though would require dedicated algorithms to be put in place before the start of the HL-LHC to achieve the full potential. Figure 1 also shows the limits for searches that have a non-trivial background component. While many of the models described in the introduction are nearly background-free, some may have contributions from long tails of the Standard Model. With a signal-to-background ratio of 0.001, a cross section limit of about 1 pb.

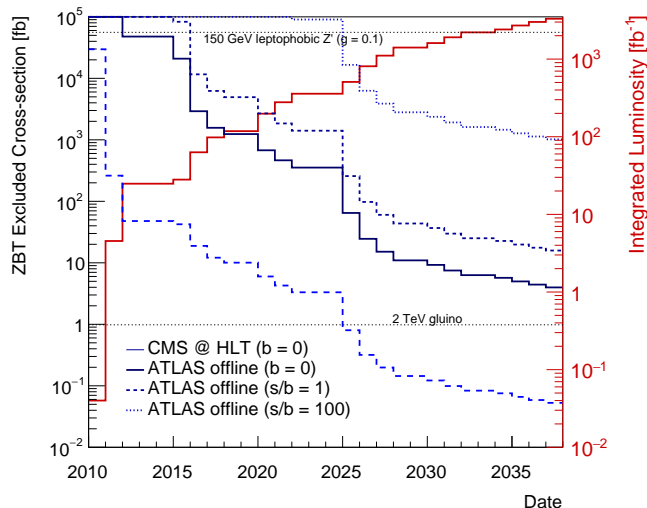


FIG. 1: The excluded cross-section and the integrated luminosity are shown throughout the lifetime of the (HL-)LHC. Statistical uncertainties are assumed to dominate over systematic ones in this low background regime. For reference, the cross-section for a 2 TeV gluino and a 150 GeV leptophobic Z' with $g = 0.1$ are also shown (see text for details).

In addition to low background searches for exotic signatures, the ZBT can also be competitive with searches for SM-like final states that exploit associated production. One important example is the low mass dijet res-

onance search. Every hadron collider has searches for dijet resonances, which are predicted in a wide variety of BSM models. The SM dijet cross-section is so large that a traditional search in the dijet invariant mass (m_{jj}) spectrum is plagued by large prescales at low m_{jj} . Such searches are therefore performed with trigger-level and ISR analyses. Figure 2 shows the effective ZBT prescale compared with the effective prescale due to the reduction in cross section when requiring the Z' to be produced in association with a ISR high p_T photon or jet. By construction, the prescale is independent of p_T for the ZBT, but grows quickly with the photon or jet p_T for the ISR analyses. Typical minimum requirements for unprescaled single photon and jet triggers are $p_T = 100$ and $p_T = 400$ GeV respectively. At these values the effective prescale is significantly larger than the one expected for the ZBT. The photon and jet thresholds can be lowered when combining the ISR technique with data scouting. However, when the ZBT is combined with data scouting (at the HL-LHC), the solid line in Fig. 2 becomes the dashed one. Therefore, even at trigger-level, the ZBT has a lower effective prescale than ISR techniques.

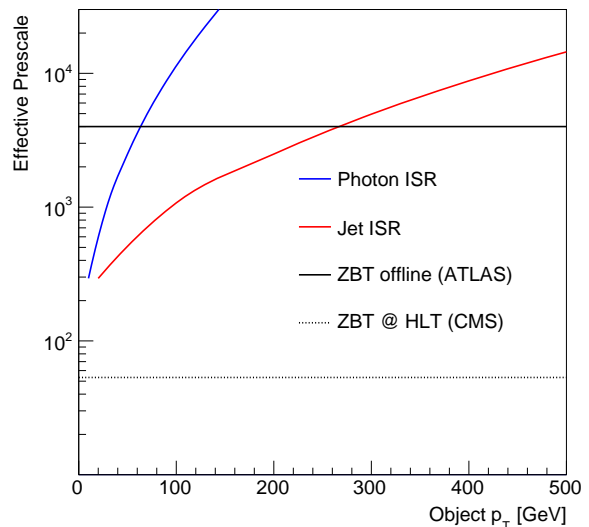


FIG. 2: A comparison of the effective prescale for the ZBT and ISR analyses for a particular leptophobic Z' model [14]. The ZBT prescales are the same as in Table I. Prescales for the ISR analyses are computed with MG5.aMC 2.1.1 [25] by comparing the cross-section for $p p > Z' j$ or $p p > Z' a$ to $p p > Z'$ and do not include any further prescale obtained at L1 or HLT.

The ZBT is superior to the ISR technique in terms of prescale (at the HL-LHC), but a fair comparison also requires an assessment of the signal-to-background ratio. A loss in events from the effective prescale from an ISR requirement can be partially compensated by better discrimination power. To estimate the approxi-

mate sensitivity to a dijet resonance, a benchmark Z' model [14] is simulated with MG5_aMC 2.1.1 [25] interfaced with Pythia 8.170 [26]. To simulate the detector response, the jet momenta are smeared according to $\sigma(p_T)/p_T = 1.3/\sqrt{p_T/\text{GeV}}$. The jet resolution depends on the pileup conditions and is in general worse at trigger-level than for fully reconstructed offline jets. Therefore, the resolution function is conservatively chosen to be worse than the typical energy resolution at LHC experiments in Run 2. Events are required to have exactly two jets with $p_T > 25$ GeV, and the two leading such jets are used to compute the dijet invariant mass, m_{jj} . More sophisticated approaches could better exploit events with significant initial or final state radiation for an enhanced sensitivity but are beyond the scope of this paper. A simple binned χ^2 analysis of the dijet invariant mass spectrum in a window around the target Z' mass is performed, using toys to estimate the p -value. A given mass point is declared excluded if the corresponding p -value is less than 0.05. As a validation of this procedure, the coupling upper limit is estimated for a 500 GeV Z' with 20.3 fb^{-1} of unprescaled single jet trigger simulated data at $\sqrt{s} = 8$ TeV. The limit obtained, approximately 1.5, is consistent with the Run 1 ATLAS result [27].

Figure 3 shows a comparison between published ISR limits and our estimate of the ZBT @ HLT based on the standard search for a peak in the m_{jj} distribution. The ISR result will slowly improve with more integrated luminosity so for a fair comparison, both strategies are evaluated with a dataset size corresponding to the 2015 run. The ZBT results are estimated assuming that the same dataset is recorded at HL-LHC rates so a relative comparison between strategies near their peak performance is possible. With this setup, the limits are found to be comparable. Given the complementarity of the two strategies, further gain can be achieved by combining the results. Note that a conservative estimate of the jet resolution at the HL-LHC is used for the ZBT analysis (which is also highly simplified). It is likely that the limits shown here are therefore conservative for the ZBT (also supported by Fig. 2).

IV. CONCLUSIONS

The multiple pileup interactions produced in LHC collisions yield unbiased data which can be used to probe physics processes otherwise inaccessible or with limited acceptance. The effective prescale for this Zero Bias Trigger is about 40000 in Runs 2+3, and drops to 400 for trigger-level analyses. If the trigger efficiency for any search is lower than this amount, then the ZBT is more powerful. In particular, for exotic signatures that are nearly impossible to trigger on due to bandwidth and time constraints in the trigger, the ZBT may be the best strategy. This includes models that have not yet appeared in the literature. For existing models, one can fully exploit the ZBT by implementing selections in the

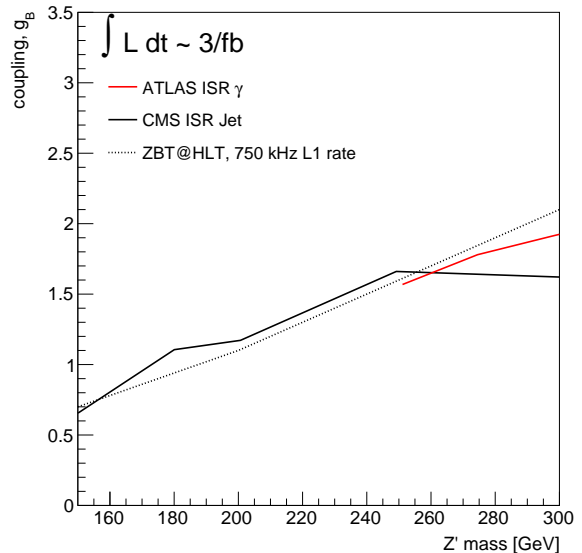


FIG. 3: The coupling limits as a function of the Z' mass for ISR jet/photon analyses using the 2015 LHC datasets compared with estimates for the ZBT using a dataset of the same size as collected at the HL-LHC (with a significantly higher bandwidth). Existing limits are from Ref. [16].

software trigger to set the most stringent limits despite failing a direct hardware trigger.

In addition to setting a bottom line for searches with a low background rate, the ZBT can also be competitive with traditional searches that exploit associative production to pass the trigger. Requiring a second object, such as a high p_T ISR jet or photon introduces a large effective prescale that can be harsher than the prescale from the ZBT. When combined with a trigger-level analysis, the ZBT is expected to provide comparable limits to the ISR technique in the case of the low mass dijet resonance search. These estimates are based on a simplified model of the dijet resonance searches and could be improved with additional sophistication. The simple model ignores the efficiency for reconstructing primary vertices and any inefficiencies in associating jets to these vertices. For the relatively high masses targeted by the example, these inefficiencies are relatively small. However, for lower mass measurements and searches, these inefficiencies may be important. Both ATLAS and CMS have ongoing studies to improve the tracking performance in high pileup environments, including the use of timing information to distinguish objects from spatially overlapping vertices. All of these interesting developments will be important for the ZBT strategy.

After the full LHC program, the ZBT will have accumulated about $0.5\text{--}1 \text{ fb}^{-1}$ of fully unbiased pp collision data. Up until now, these data would have been analyzed. We have shown that the novel concept of analyzing all pileup interactions enhances the physics reach of

the LHC experiments and will constitute an important strategy to fully exploit the HL-LHC dataset.

V. ACKNOWLEDGMENTS

We would like to thank Felix Yu for help simulating the Z' benchmark model and Nhan Tran for useful dis-

cussions about the ZBT. In addition, we are grateful to John Alison for helping us compute the correct luminosity for the ZBT.

-
- [1] **ATLAS** Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, 1606.02625.
 - [2] **CMS** Collaboration, *Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV*, *CMS-PAS-FSQ-15-005* (2016).
 - [3] **ATLAS** Collaboration, *Measurement of W^\pm and Z -boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **759** (2016) 601–621 [1603.09222].
 - [4] **CMS** Collaboration, *Measurement of inclusive W and Z boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV*, *CMS-PAS-SMP-15-004* (2015).
 - [5] **ATLAS** Collaboration, *Measurement of jet charge in dijet events from $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector*, *Phys. Rev. D* **93** (2016) 052003 [1509.05190].
 - [6] P. Meade, M. Papucci and T. Volansky, *Odd Tracks at Hadron Colliders*, *Phys. Rev. Lett.* **109** (2012) 031801 [1103.3016].
 - [7] L. B. Okun, *THETONS*, *JETP Lett.* **31** (1980) 144–147. [Pisma Zh. Eksp. Teor. Fiz.31,156(1979)].
 - [8] L. B. Okun, *THETA PARTICLES*, *Nucl. Phys. B* **173** (1980) 1–12.
 - [9] J. Kang and M. A. Luty, *Macroscopic Strings and 'Quirks' at Colliders*, *JHEP* **11** (2009) 065 [0805.4642].
 - [10] S. Knapen, S. Pagan Griso, M. Papucci and D. J. Robinson, *Triggering Soft Bombs at the LHC*, 1612.00850.
 - [11] R. Harnik and T. Wizansky, *Signals of New Physics in the Underlying Event*, *Phys. Rev. D* **80** (2009) 075015 [0810.3948].
 - [12] P. Schwaller, D. Stolarski and A. Weiler, *Emerging Jets*, *JHEP* **05** (2015) 059 [1502.05409].
 - [13] T. Cohen, M. Lisanti and H. K. Lou, *Semivisible Jets: Dark Matter Undercover at the LHC*, *Phys. Rev. Lett.* **115** (2015), no. 17 171804 [1503.00009].
 - [14] B. A. Dobrescu and F. Yu, *Coupling-mass mapping of dijet peak searches*, *Phys. Rev. D* **88** (2013) 035021 [1306.2629].
 - [15] **ATLAS** Collaboration, *Search for new light resonances decaying to jet pairs and produced in association with a photon in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *ATLAS-CONF-2016-029* (2016).
 - [16] **CMS** Collaboration, *Search for light vector resonances decaying to quarks at 13 TeV*, *CMS-PAS-EXO-16-030* (2016).
 - [17] **CMS** Collaboration, *Search for narrow resonances in dijet final states at $\sqrt{s} = 8$ TeV with the novel CMS technique of data scouting*, *Phys. Rev. Lett.* **117** (2016) 031802 [1604.08907].
 - [18] **ATLAS** Collaboration, *Search for light dijet resonances with the ATLAS detector using a Trigger-Level Analysis in LHC pp collisions at $\sqrt{s} = 13$ TeV*, .
 - [19] **ATLAS** Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, 1510.03823.
 - [20] **ATLAS** Collaboration, *ATLAS Phase-II Upgrade Scoping Document*, *CERN-LHCC-2015-020* (2015), no. CERN-LHCC-2015-020. LHCC-G-166.
 - [21] **CMS** Collaboration, *CMS Phase II Upgrade Scope Document*, *CERN-LHCC-2015-019* (2015), no. CERN-LHCC-2015-019. LHCC-G-165.
 - [22] A. L. Read, *Presentation of search results: The $CL(s)$ technique*, *J. Phys. G* **28** (2002) 2693–2704. [,11(2002)].
 - [23] C. Borschensky, M. Kramer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn and X. Portell, *Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13, 14, 33$ and 100 TeV*, *Eur. Phys. J. C* **74** (2014), no. 12 3174 [1407.5066].
 - [24] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930 [1112.5675].
 - [25] J. Alwall et. al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [1405.0301].
 - [26] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867 [0710.3820].
 - [27] **ATLAS** Collaboration, *Search for new phenomena in the dijet mass distribution using pp collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Phys. Rev. D* **91** (2015) 052007 [1407.1376].